

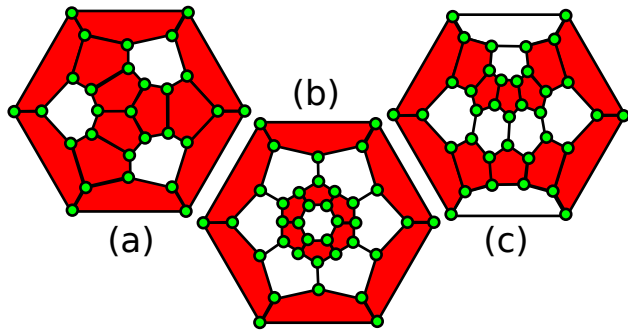
# Structures and pathways for clathrin self-assembly in the bulk and on membranes

Richard Matthews<sup>a\*</sup> and Christos N. Likos<sup>a‡</sup>

We present a coarse-grained model of clathrin that is simple enough to be computationally tractable yet includes key observed qualitative features: a triskelion structure with excluded volume between legs; assembly of polymorphic cages in the bulk; formation of buds on a membrane. We investigate the assembly of our model using both Monte Carlo simulations and molecular dynamics with hydrodynamic interactions, in the latter employing a new membrane boundary condition. In the bulk, a range of known clathrin structures are assembled. A membrane budding pathway involving the coalescence of multiple small clusters is identified.

## 1 Introduction

Clathrin<sup>1</sup> is a triskelion-shaped protein that self-assembles into a broad range of polymorphic structures. On the one hand, its key function is in forming coated vesicles, separated from membranes through budding, that are crucial for intra-cellular transport<sup>2</sup>. On the other hand, its three-legged shape lets it also form extended hexagonal sheets<sup>3</sup>. *In vivo*, clathrin assembly is always associated with membranes: it is attached to them by intermediary protein complexes called adaptors.



**Fig. 1** Schlegel diagrams of different structures assembled by clathrin, based on similar diagrams in ref.<sup>4</sup>. (a) Mini-coat. (b) Hexagonal barrel. (c) Tennis ball. The green dots represent the centres of triskelia and the black lines joining them represent the connections between them, formed by legs lying next to each other and bonding. The red areas show pentagonal faces and the white ones show hexagonal faces. Note that the projection of the three-dimensional cage structures onto the plane does not preserve relative lengths or angles. For each diagram, the surrounding white space represents an additional hexagonal face in the three-dimensional structure.

Assembly of cages may also be observed in *in vitro* experiments without a membrane: early work<sup>5</sup> found that the resulting cages were much more homogeneous when adaptor proteins were present. A number of closed-cage structures have been identified, all having twelve pentagonal faces and  $(N - 20)/2$  hexagonal faces, where  $N$  is the number of triskelia. These include one structure with  $N = 28$ , called a mini-coat, two with  $N = 36$ , given the names hexagonal barrel and tennis ball, and a truncated icosahedron with  $N = 60$ <sup>4,6,7</sup>. Of these, the tennis ball structure, with a closed ring of pentagons reminiscent of the seam on a tennis ball, is less common. The mini-coat, tennis ball and hexagonal barrel structures are illustrated in Fig. 1. Larger cages may also be formed. Detailed investigation of coated vesicles showed them to be much more poly-disperse, including some heptagons, although the tennis ball structure was also observed<sup>8</sup>.

Previous modeling of the assembly of structures by clathrin<sup>9,10</sup> assumed that the triskelia are completely rigid. However, there is evidence, based on analysis of the fluctuations observed in electron micrograph images<sup>11</sup> and the comparison of Brownian dynamics simulations to scattering data<sup>12</sup>, that, in isolation, the legs of the triskelion have a persistence length similar to their contour length  $\approx 50\text{nm}$ <sup>1</sup>. It is however both expected<sup>11</sup> and observed<sup>13</sup> that there is much greater rigidity once the triskelia are bonded into a structure.

Each triskelion leg is primarily composed of an extended, curved sub-unit called a heavy chain<sup>1</sup>. Much of the internal construction of the heavy chain comprises zig-zag structures. The leg flexibility within a cage was estimated, through observation of crystal structures, to allow bends of  $1^\circ - 2^\circ$  per zig-zag<sup>13</sup>. This estimate may be too high for the leg overall as it was based on a section known as the linker, which is expected to be more flexible due to a less regular structure<sup>14</sup>.

In this work, we present a new clathrin model that includes excluded volume. Each leg is modeled by a sequence of bonded patchy particles. The interactions between patchy particles typically have strong orientational dependence. In com-

<sup>a</sup> Faculty of Physics, University of Vienna, Boltzmannngasse 5, A-1090 Vienna, Austria.

\* E-mail: [richard.matthews@univie.ac.at](mailto:richard.matthews@univie.ac.at)

‡ E-mail: [christos.likos@univie.ac.at](mailto:christos.likos@univie.ac.at)

putational models, this may be included directly in pair interaction potentials<sup>15</sup> or be produced by composing sub-units of multiple particles. This latter approach may give a more realistic representation of the shape of sub-units and has been applied to viral capsids, both in studies of self-assembly<sup>16</sup> and in modelling interactions with membranes without assembly<sup>17</sup>, where budding was observed. Viral capsids are perhaps the most intensively studied example of self-assembly and models with single particle sub-units have also been applied<sup>18</sup>. For the case of clathrin, whilst previous work<sup>9,10</sup> has also used such simpler models, here our approach is intermediate: sub-units composed of multiple particles whose pair interactions are patchy.

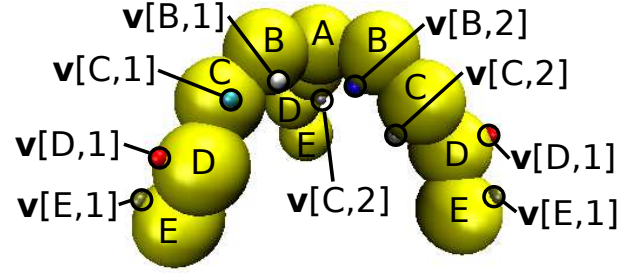
Although the exact form of the attractive interaction between clathrin legs is not known<sup>9</sup>, in observed structures<sup>1</sup> they tend to lie close to each other, always having a similar relative orientation, suggesting interactions are short range and strongly orientationally dependent. Whilst, particularly for viral capsids, the use of patchy particles to represent protein-protein interactions is quite common<sup>15,18,19</sup>, we furthermore choose to employ them as an efficient way to capture the two key interaction features: short range and strong orientational dependence.

In a recent publication<sup>15</sup>, we presented results on the effect of fluctuating membranes on the equilibrium structures of a system of self-assembling patchy colloids. We considered a simple model, representing each clathrin with a single spherical particle with three attractive patches. Here, our use of multiple patchy beads allows, in contrast to previous approaches<sup>9,10</sup>, features that are expected to be important in self-assembly to be captured: excluded volume between legs, flexibility, the interweaving of legs in assembled structures.

The rest of the paper is organised as follows. In Sec. 2, we describe our model in more detail, including the process used to determine parameters. In Sec. 3, we present the results of Monte Carlo (MC)<sup>20</sup> simulations to explore the structures that our triskelia may assemble in the bulk, before moving to dynamical simulations to consider the behaviour with a membrane in Sec. 4. In Sec. 5, we summarise and draw our conclusions, whereas in the Appendix we present some technical details pertaining to the model and the simulation techniques.

## 2 The clathrin model

Our model triskelion comprises 13 bead patchy beads, see appendix A.1, of 5 different types, denoted  $\alpha = A - E$ . One central bead of type  $A$  is attached to 3 legs, each consisting of 1 bead each of types  $B - E$ , see Fig. 2. Associated with bead type  $\alpha$  are two sets of unit vectors:  $\{\mathbf{v}[\alpha, i]\}$  and  $\{\mathbf{u}[\alpha, i]\}$ , where  $i$  indexes the different vectors belonging to one type.  $\{\mathbf{v}[\alpha, i]\}$  specify the attractive patches for interactions with other triskelia, whereas  $\{\mathbf{u}[\alpha, i]\}$  define the internal interactions



**Fig. 2** Depiction of a triskelion comprised of beads type  $\alpha = A - E$ , with patches whose positions are defined by  $\{\mathbf{v}[\alpha, i]\}$ . There are attractive interactions between different triskelia, between patches  $\mathbf{v}[B, 1]$  and  $\mathbf{v}[C, 1]$ , patches  $\mathbf{v}[B, 2]$  and  $\mathbf{v}[D, 1]$  and patches  $\mathbf{v}[C, 2]$  and  $\mathbf{v}[E, 1]$ . Note that all legs are identical.

tions and thus the shape of an isolated triskelion in mechanical equilibrium. A further parameter,  $d$ , specifies the mechanical equilibrium separation between bonded beads of the same triskelion. A detailed description of how the parameters determine the triskelion shape is given in appendix A.2.

Considering  $\{\mathbf{v}[\alpha, i]\}$ , apart from type  $A$ , all beads have attractive patches, depicted in Fig. 2, which are supplemented by a torsional vector (not shown). Moving out from the centre, type  $B$  has two patches, whose positions are defined by  $\mathbf{v}[B, 1]$  and  $\mathbf{v}[B, 2]$ , as does type  $C$ . Patches  $\mathbf{v}[B, 1]$  and  $\mathbf{v}[C, 1]$ , from different triskelia, attract each other such that, if the initial parts of two legs are placed approximately antiparallel, and at an appropriate separation, they may bond. The other patches of  $B$  and  $C$ ,  $\mathbf{v}[B, 2]$  and  $\mathbf{v}[C, 2]$ , attract the single patches on types  $D$  and  $E$ ,  $\mathbf{v}[D, 1]$  and  $\mathbf{v}[E, 1]$  respectively. These interactions are such that, if triskelia form a cage, the second half of a given leg bends under the first half of a leg from an adjacent triskelion, whereby “under” means towards the centre of the cage, mimicking nature<sup>1</sup>. Our choice of patches for the different beads, and of the specificity of their interactions, is made so that the parts of the legs that are observed to lie next to each other in clathrin structures<sup>1</sup> will attract each other in our model.

There are excluded volume interactions between all beads not belonging to the same triskelion. The shape of the triskelion is maintained by internal interactions: harmonic springs with equilibrium length  $d$  and spring constant  $k$  between bonded beads, plus bending and torsional stiffness. We choose the bending and torsional rigidity to be the same, specified by the parameter  $\kappa$ . The equilibrium angles between subsequent bonds along a leg are encoded in the  $\{\mathbf{u}[\alpha, i]\}$  vectors. Full definition of the interaction potentials for both external and internal interactions is given in the appendices A.1 and A.2 respectively. The extended nature of the legs, and the specificity of the attractive patches, gives the overall interaction be-

tween two triskelia orientational dependence. This is further enforced at a bead-bead interaction level through the torsional vectors, see appendix A.1.

The parameters of triskelion shape and patch vectors are not set *a priori* but they are rather specified through an informed search procedure. The basic idea of our parameter-finding scheme is the following: hold a set of objects in a desired configuration, which here will be the hexagonal barrel comprising 36 triskelia, and allow the parameters determining the interactions between these objects to vary until they have found a low energy minimum. We assume that if the interaction parameters are then fixed, they will drive the objects to reform the structure from a random initial condition. We use the term “free assembly” to refer to simulations where interaction parameters are fixed. For free assembly simulations we consider both random and pre-assembled initial conditions. Whilst in the latter, strong interactions may cause to triskelia to remain in their initial configuration, unlike in the parameter-finding, triskelia may in principle explore other structures.

The parameters for triskelion shape  $\{\mathbf{u}[\alpha, i]\}$ , patch position  $\{\mathbf{v}[\alpha, i]\}$ , and bead separation  $d$  were chosen using Metropolis MC simulations, in which these parameters, along with the usual system coordinates, were treated as dynamical variables. Updates were made using trial moves with the standard acceptance criterion<sup>20</sup>. Schematically, a Hamiltonian  $H^*(\mathbf{X}, d^*, \mathbf{V}^*, \mathbf{U}^*)$  was used to determine the parameters, which were then input to the Hamiltonian  $H(\mathbf{X}; d, \mathbf{V}, \mathbf{U})$  for free assembly simulations.  $\mathbf{V}$  and  $\mathbf{U}$  represent the set of all parameter vectors for all bead types and  $\mathbf{X}$  represents the usual system coordinates. For  $H$ , the values of  $d$ ,  $\mathbf{V}$  and  $\mathbf{U}$  are fixed. The corresponding variables in  $H^*$ ,  $d^*$ ,  $\mathbf{V}^*$  and  $\mathbf{U}^*$  may vary freely. Simulations with  $H^*(\mathbf{X}, d^*, \mathbf{V}^*, \mathbf{U}^*)$  were performed at low temperature,  $k_B T \ll \epsilon_{tt}$ , where  $-\epsilon_{tt}$  is the minimum of the attractive interaction between beads, so that the system relaxed to a low-energy minimum. Here  $k_B T$  is the energy appearing in the standard Metropolis Monte Carlo acceptance probability<sup>20</sup>,  $\min[1, \exp(-\Delta E/k_B T)]$ , where  $\Delta E$  is the change in the energy due to a trial move. Additional constraints were applied to ensure the minimum found corresponded to the desired structure. To extract parameter values for use in simulations with  $H(\mathbf{X}; d, \mathbf{V}, \mathbf{U})$ , thermal averages of the corresponding variables around the minimum were performed. To simplify the minimisation, during the interaction-finding stage, all triskelia always had a configuration corresponding to the minimum of their internal interactions, see appendix A.3 for more details.

A common self-assembled shape observed in *in vitro* experiments with clathrin is the hexagonal barrel<sup>7</sup> and we chose this as our target structure. Clearly, a different choice of target structure would lead to a somewhat different set of interaction parameters but, given its frequency in bulk assembly experiments, the hexagonal barrel is a reasonable choice. We fur-

thermore emphasise, however, that, in free assembly runs (see Sec. 3), our triskelia were also able to self-assemble into different structures. Multiple parameter-finding runs at different temperatures were observed to give very similar parameters, see appendix A.3. The set of parameters used in our free assembly simulations is given in appendix A.3. The state found cannot be guaranteed to correspond to the global minimum for a hexagonal barrel. However, given the tightness of the packing of the triskelia observed in the final structure, it is a reasonable assumption that the configuration is the unique minimum for triskelia interacting in the desired way, with initial parts of legs lying side-by-side and antiparallel, and the end parts of legs tucked inside the cage.

We set the parameter for the harmonic springs between beads, which is not varied, to  $k = 1.6 \times 10^3 k_B T$ , and consider different bending stiffnesses,  $\kappa$  and patch attraction strengths,  $\epsilon_{tt}$ . Our simple model represents only those two sections of a leg, which when assembled run along two polyhedron edges. In the corresponding section of a true triskelion leg there are  $\approx 20$  zig-zags<sup>1</sup>. We primarily consider bending stiffness parameters of  $\kappa = 0.8 \times 10^3 k_B T$ ,  $1.6 \times 10^3 k_B T$  and  $3.2 \times 10^3 k_B T$ .  $\kappa = 0.8 \times 10^3 k_B T$  gives a typical angular deflection per bending joint of  $\approx 3^\circ$ . Since there are eight joints in a leg, one at each end of each internal bond, see appendix A.2, this gives a total possible deflection per leg similar to that expected from crystal structure observations<sup>1,13</sup>. We also considered complete rigidity, applied also to the springs joining beads, as well as flexibility similar to that seen for isolated triskelia. In the latter case we found assembly of disordered and extended structures rather than cages and results are not presented. It should be noted, however, that the structures found for stiffer triskelia will also represent local minima for the flexible ones, although in this case our free assembly simulations were unable to find them.

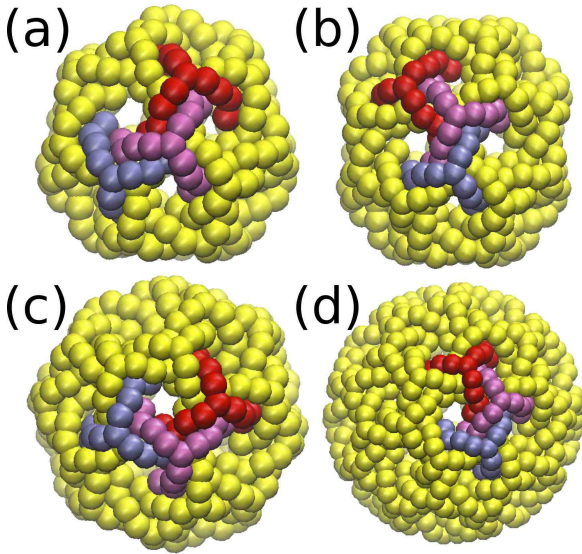
### 3 Bulk self-assembly

We next consider the structures formed by our clathrin-model without a membrane. For these free assembly simulations, we employ Metropolis Monte Carlo with a range of moves to improve sampling, including Aggregate Volume Bias<sup>21</sup>, Configurational Bias<sup>22</sup>, cluster moves<sup>23</sup>, Hybrid MC<sup>24</sup> and multi-canonical parallel tempering<sup>25</sup>. To form closed cages, triskelia must be able to bond and form faces surrounded by both 5- and 6-edge loops. This flexibility, which is automatically built into our model through the parameter choosing procedure, means that the triskelia may explore a broad range of competing low-energy minima. Whilst we expect the global minimum to be a closed cage, simulations may easily become trapped in other states and, despite the range of MC moves utilised, we find that simulations are not able to move between all of the local minima on a feasible timescale. Nonetheless, the free as-



sembly simulations do give us reliable information about the structures that our model may assemble.

We ran free assembly simulations with  $N = 36$  triskelia, starting from two initial configurations: one with triskelia placed randomly, just with the requirement of no beads overlapping, and the other with an assembled hexagonal barrel. 24 systems with different  $\epsilon_{tt}$  between  $3.3k_B T$  and  $4.91k_B T$  were run with parallel tempering swaps between them. In the majority of simulations, umbrella sampling with an iteratively-calculated weighting function<sup>25</sup>,  $w(U_{tt}/\epsilon_{tt})$ , where  $U_{tt}$  is the total inter-triskelion interaction energy, was used but some runs were also performed without. We first, in Fig. 3, present results from individual free assembly simulations for various quantities as a function of  $\epsilon_{tt}$ .



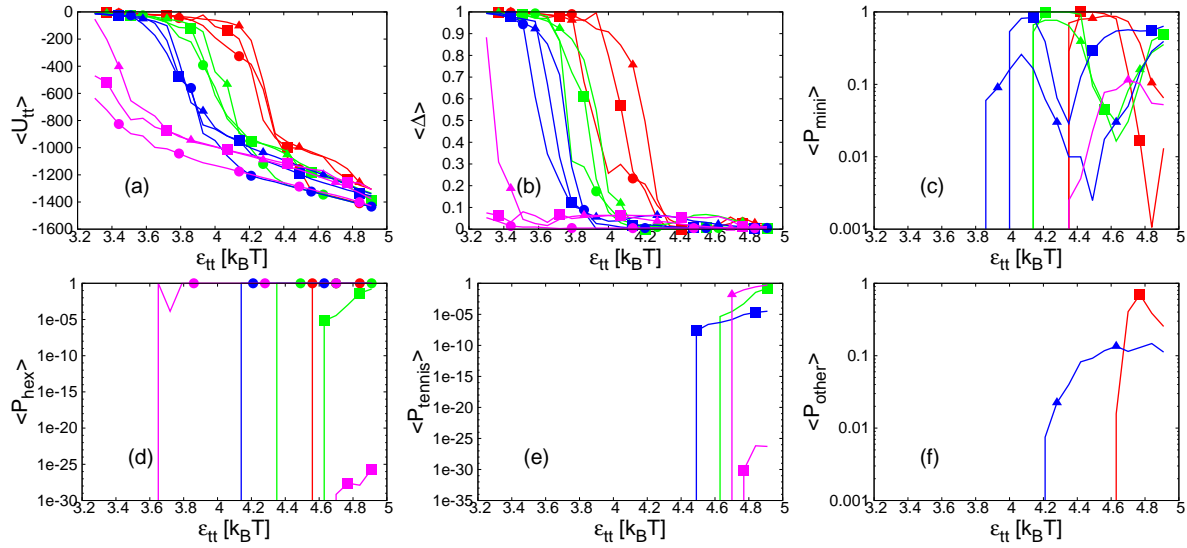
**Fig. 4** Assembled structures: (a) Mini-coat with  $N = 28$  triskelia (b) Hexagonal barrel,  $N = 36$ . (c) Tennis ball,  $N = 36$ . (d) Truncated icosahedron,  $N = 60$ . In each, three triskelia are shown in different colours to highlight their relative positions. The first three snapshots are all from free assembly simulations. The truncated icosahedron was not observed in free assembly so for the final configuration triskelia were placed by hand. The structure depicted is nonetheless mechanically stable for sufficient interaction strengths, see text and Fig. 5.

Fig. 3(a) shows the average inter-triskelion interaction energy,  $\langle U_{tt} \rangle$ . For low  $\epsilon_{tt}$ , the results, at least for flexible triskelia, are similar for both assembled and unassembled initial configurations. However, for higher  $\epsilon_{tt}$ , the results for the assembled initial condition clearly show lower energies, indicating that proper sampling of equilibrium is not achieved. The point at which assembly starts is at higher  $\epsilon_{tt}$  for more flexible triskelia, due to the larger loss of entropy. For the asphericity<sup>26</sup> of the largest cluster,  $\langle \Delta \rangle$ , broadly similar results

are seen for both initial conditions, with  $\langle \Delta \rangle$  changing from  $\approx 1$  indicating highly aspherical structures when there is little assembly to  $\approx 0$  indicating almost spherical structures at high  $\epsilon_{tt}$ . In calculating the asphericity, the positions of the central, type *A* beads were used. When there are no bonded triskelia in the system and thus the largest “cluster” is a single triskelion, the asphericity, which is calculated from a tensor based on the separations of pairs of beads<sup>26</sup>, is not defined. Since, when the largest cluster is of size 2, which is necessarily a line,  $\Delta = 1$ , we choose to assign a value of  $\Delta = 1$  for a single triskelion also.

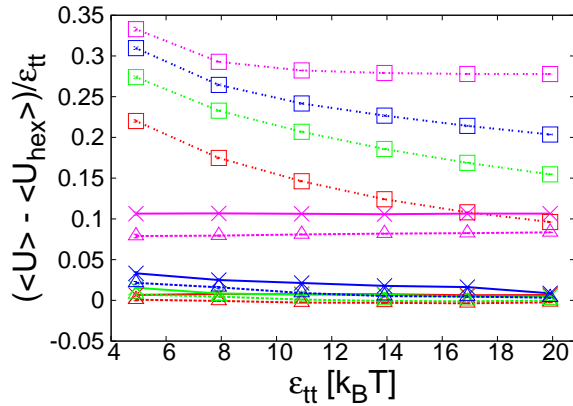
Much larger differences are seen in Fig. 3(c) - (f), where the probability of observing specific structures is considered. Here we plot the probabilities on logarithmic scales, down to very small values. These very low probabilities arise from the umbrella sampling: during the creation of the weighting function the system may become trapped in some configuration, eventually the weighting function will become large enough to allow the system to escape and explore other structures. However, due to the large weighting function, the estimated probability of these structures is very low. Since we know our simulations are not fully sampling equilibrium, these probabilities may well be severely underestimated. We check whether a bonded cluster has one of the four common structures - mini-coat, hexagonal barrel, tennis ball or truncated icosahedron - by using an algorithm<sup>27</sup> to test if the graph formed by considering the bonds between triskelia is isomorphic to the one of the corresponding graphs. We identify two beads from different triskelia as being bonded if their interaction energy is  $< -\frac{1}{4}\epsilon_{tt}$ . Two triskelia are then defined to be bonded if there exists at least one bond between their type *B* and type *C* beads.

For the unassembled initial condition, we find that, overall, the most likely structure to be formed is the mini-coat, see Fig. 3(c), both for simulations with and without umbrella sampling. At higher values of  $\epsilon_{tt}$ , some initially unassembled simulations did also form hexagonal barrel structures, consistent with the fact that interactions were chosen for this structure at low temperature, and also tennis ball structures, see Fig. 3(d) and (e). In Fig. 3(f) we show the probability of forming closed structures that have twelve pentagonal faces and  $(N - 20)/2$  hexagonal faces, but which are not one of the known structures that we test for. These other structures that arose in our simulations had 32 triskelia. No closed structures with a different number of pentagonal and hexagonal faces were formed, although there were additionally many open structures. For the assembled initial condition we found that the only closed structure seen in the simulation was the hexagonal barrel. We also ran free assembly simulations with 60 triskelia but no truncated icosahedra were assembled although, when pre-assembled, they were stable for higher  $\epsilon_{tt}$ . In Fig. 4 we show snapshots of mini-coat, hexagonal barrel and tennis ball



**Fig. 3** Average values as a function of  $\epsilon_{tt}$  from single MC free assembly simulations with different flexibilities:  $\kappa = 0.8 \times 10^3 k_B T$  (red),  $\kappa = 1.6 \times 10^3 k_B T$  (green),  $\kappa = 3.2 \times 10^3 k_B T$  (blue), rigid (magenta). Different simulation types: unassembled initial condition with umbrella sampling (■); unassembled initial condition without umbrella sampling (▲); assembled initial condition with umbrella sampling (●). (a) Total inter-triskelion interaction energy. (b) Asphericity of the largest cluster in the system. (c) Mini-coat probability. (d) Hexagonal barrel probability. (e) Tennis ball probability. (f) Probability of different closed structure with the expected number of pentagons and hexagons, see text.

structures assembled in our simulations, as well as a truncated icosahedron structure.



**Fig. 5** Difference between average internal energy per triskelion of a given structure - mini-coat (×, solid line), tennis ball (△, dashed line) or truncated icosahedron (□, dot-dashed line) - and that of hexagonal barrel for different bending rigidities,  $\kappa = 0.8 \times 10^3 k_B T$  (red),  $\kappa = 1.6 \times 10^3 k_B T$  (green),  $\kappa = 3.2 \times 10^3 k_B T$  (blue) and rigid (magenta). Note that for the lowest  $\epsilon_{tt}$  for rigid triskelia, the mini-coat structure was unstable and disassembled: the data point plotted is only averaged over those parts of the simulations before disassembly occurred.

To obtain more information about the relative stability of the different structures, also for higher  $\epsilon_{tt}$ , we consider, in Fig. 5, the average internal energy of a given structure divided by the number of triskelia in the structure, compared to the value for a hexagonal barrel. Simulations were run at single  $\epsilon_{tt}$  values from  $4.91 k_B T$  to  $19.91 k_B T$  with only local moves. The initial condition was taken as the assembled structure and for all simulations, except in one case, this structure persisted for the rest of the simulation: for the lowest  $\epsilon_{tt}$  the structure was intermittently not identified according to our bonding definition, though only temporarily, indicating that true disassembly had not occurred. For rigid triskelia, the mini-coat with  $\epsilon_{tt} = 4.91 k_B T$  was not stable. Although the overall structure did not disassemble, typically multiple bonds within the structure broke and did not reform within the simulation. It should be noted, however, that, given the mini-coat did persist for some time, and also given it was formed in some free assembly simulations with rigid triskelia, see Fig. 3(c), at  $\epsilon_{tt} = 4.91 k_B T$  the mini-coat must still represent a local minimum for the rigid triskelia.

We find that, usually, the hexagonal barrel has the lowest internal energy per triskelion, as expected since the parameters were determined for this structure. We find that the difference becomes more positive as the rigidity is increased but, for the highest two flexibilities, the tennis ball has lower internal energy per triskelion than the hexagonal barrel at some  $\epsilon_{tt}$ . The

mini-coat has values relatively close to those for the hexagonal barrel but the truncated icosahedron has significantly higher values for all  $\epsilon_{tt}$  considered. We observe that the difference relative to  $\epsilon_{tt}$  decreases as  $\epsilon_{tt}$  increases since the attractions dominate more over the bending rigidity but that for completely rigid triskelia the curves flatten out.

#### 4 Self-assembly on a membrane: hydrodynamics and bud formation

Clathrin is intrinsically linked to membranes, and *in vivo* it is here that its self-assembly occurs<sup>2</sup>. As discussed in Sec. 1, the structures formed on a membrane are more poly-disperse than those assembled in the bulk and it is expected that the fluctuating surface will also change the pathways to assembly. In this section, we consider the assembly of our triskelia when they are attracted to a fluctuating membrane. In the terms used in the previous section, all the simulations with a membrane in this section are free assembly ones. We represent the membrane as a dynamically triangulated surface composed of bonded particles with a typical bond length of  $\sigma$ <sup>28</sup>. Membrane fluidity is included by MC moves, performed at regular intervals, that attempt to flip bonds between neighbouring particles. The rate of bond-flipping sets the viscosity of the membrane, which may be measured by considering a Poiseuille flow in a two-dimensional membrane sheet, see appendix A.4<sup>28</sup>.

Since free assembly simulations without a membrane indicated that our MC approach is unable to fully sample equilibrium, we proceed directly to dynamical simulations. We perform molecular dynamics simulations with our triskelia, as well as the membrane, coupled to a Stochastic Rotation Dynamics (SRD) solvent<sup>29</sup>. SRD is a coarse-grained method in which the fluid is represented by point particles of mass  $m$  whose interactions are effected by dividing the system into a grid of cells at regular time intervals and exchanging momentum by a rotation through a certain angle of velocities relative to the cell centre of mass velocity. This acts as a thermostat, whilst also conserving momentum so that hydrodynamic interactions are included. More details of parameter choices and solute-solvent couplings are given in appendix A.5. Hydrodynamic interactions are naturally present in real, experimental systems and may have both qualitative and quantitative effects<sup>30</sup>. Although we do not investigate the effect of hydrodynamic interactions in detail, we choose a simulation method that includes them, as this is expected to be more dynamically realistic.

Since SRD requires the simulation box to be regularly divided into a grid with an integer number of cells, it is incompatible with the approach to simulating a tensionless membrane employed in our previous work without solvent<sup>15</sup> that involved box-rescaling. We therefore have developed a new

approach, detailed in appendix A.4, in which the edge of the membrane is bonded to a square frame, whose sides are a distance  $r_{frame}$  from the edges of the simulation box. The frame may expand and contract.

We define a unit of simulation time,  $t_0 = \sigma\sqrt{k_BT/m}$ . Our parameter choices give a membrane viscosity of  $\eta_m = 35.1 \pm 0.1m/t_0$  and a fluid viscosity of  $\eta_f = 2.5m/\sigma t_0$ , see appendices A.4 and A.5. The ratio of the viscosities is  $l_\eta = \eta_m/\eta_f \approx 14\sigma$ . For a lipid bilayer in water,  $l_\eta = 1 - 10\mu m$ <sup>31</sup>. The typical size of a triskelion is on the order of  $0.1\mu m$ , whereas in our simulation it is a few  $\sigma$ . Thus the size of a triskelion compared to  $l_\eta$  is close to the lower end of the expected range, allowing efficient simulation.

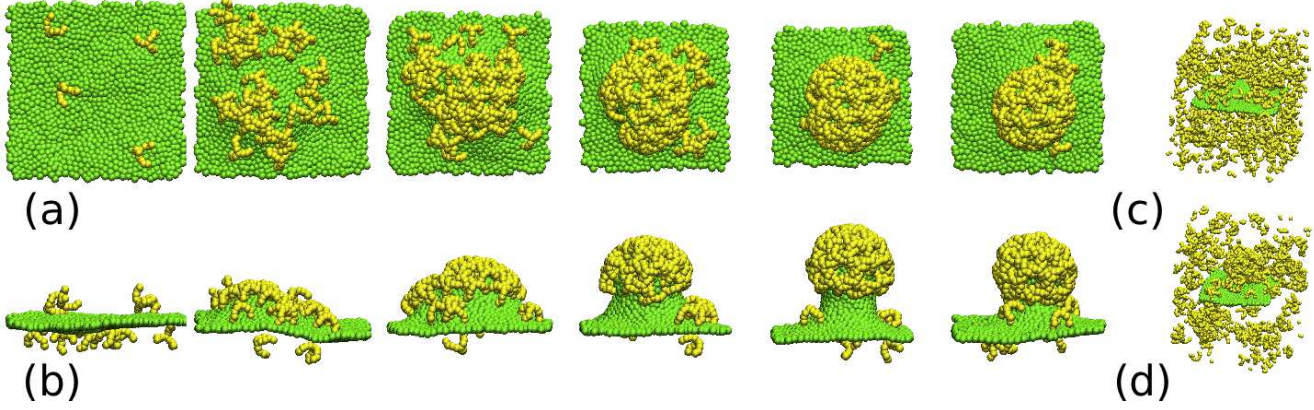
As for the MC simulations, we consider triskelion stiffnesses of  $\kappa = 0.8 \times 10^3 k_BT$ ,  $1.6 \times 10^3 k_BT$  and  $3.2 \times 10^3 k_BT$  and we consider membrane bending stiffnesses of  $\lambda_b = 2\sqrt{3}k_BT$ ,  $4\sqrt{3}k_BT$  and  $8\sqrt{3}k_BT$ . See appendix A.4 for the definition of the bending potential. Interactions of clathrin triskelia with membranes occur at the ends of the triskelion legs<sup>13</sup> via intermediary adaptor proteins. We neglect the adaptor proteins and simply introduce an attractive interaction between the final beads in the legs and membrane particles, with a minimum of  $-\epsilon_{mt}$ . Unlike for triskelion-triskelion interactions, this attraction is not patchy but the bead-bead and membrane-bead potentials share a common radial form. We consider  $\epsilon_{mt} = k_BT$  and  $2k_BT$ , and  $\epsilon_{tt} = 5k_BT$  and  $10k_BT$ .

We simulate 300 triskelia, 1156 membrane particles and  $\approx 5 \times 10^5$  SRD particles in a box of  $45\sigma \times 45\sigma \times 45\sigma$  with periodic boundaries. The relatively high triskelion density, about 10 times that used in previous work<sup>9</sup>, is chosen such that assembly proceeds quickly but we do not expect it to qualitatively affect assembly on the membrane. An equilibration period with purely repulsive interactions of  $3 \times 10^3 t_0$ , chosen to be sufficient to allow the membrane to relax, was allowed before the system was simulated with attractions for  $2.5 \times 10^4 t_0$ .

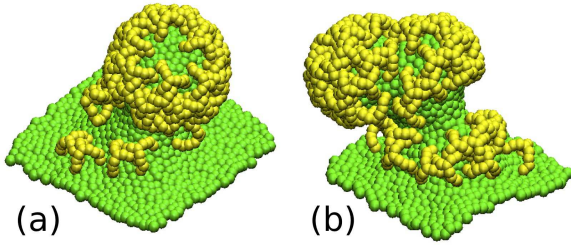
For most parameter choices, the triskelia assembled on the the membrane, causing the membrane to form a bud, see Fig. 6. The example in Fig. 6 shows the formation of a relatively defect-free, approximately spherical cage on the membrane. Often, however, the cages formed had defects or gaps of varying sizes, see Fig. 7(a). Additionally, some runs produced structures with two largely separate cages attached to one bud, causing a double-headed structure, see Fig. 7(b). Similarly lumpy structures have been observed experimentally in clathrin assembly<sup>32</sup>.

We found that  $r_{frame}$  was a good indicator of whether budding had occurred, moving to higher values as membrane area was taken into the bud and the frame contracted. In Fig. 8 we plot  $\langle r_{frame} \rangle$  against time for a variety of parameters. We first note that the results for different  $\kappa$  are similar and also that the rate of bud formation did not show strong dependence on  $\epsilon_{tt}$ . In contrast, the rate of bud formation did depend on the values





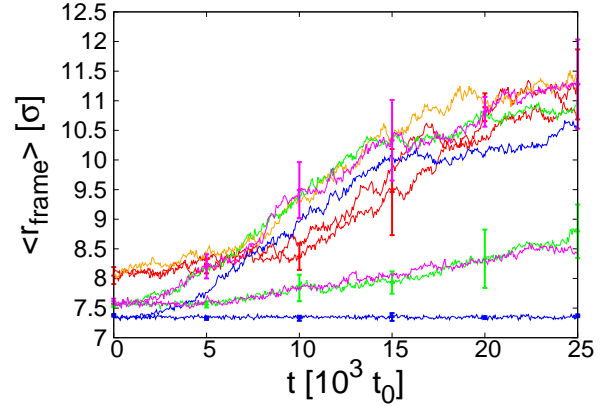
**Fig. 6** Snapshots for assembly on a membrane with  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\lambda_b = 8\sqrt{3}k_B T$ ,  $\epsilon_{tt} = 5k_B T$  and  $\epsilon_{mt} = 2k_B T$ . Triskelia are depicted in yellow and membrane particles in green. For (a) and (b) only triskelia whose central bead is within  $4\sigma$  of a membrane particle are shown: (a) Series of snapshots at regular time intervals from  $t = 0 - 2.5 \times 10^4 t_0$ . (b) Same configurations shown from different viewing angle. (c) Snapshot of same system at  $t = 0$  with all triskelia. (d) Snapshot of same system at  $t = 2.5 \times 10^4 t_0$  with all triskelia.



**Fig. 7** Snapshots of assembly on a membrane at  $t = 2.5 \times 10^4 t_0$  with  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\epsilon_{tt} = 5k_B T$  and  $\epsilon_{mt} = 2k_B T$ . Colouring as in Fig. 6, only triskelia whose central bead is within  $4\sigma$  of a membrane particle are shown. (a)  $\lambda_b = 8\sqrt{3}k_B T$ , showing an example of assembly with a gap in the cage. (b)  $\lambda_b = 4\sqrt{3}k_B T$ , showing an example of assembly with two distinct cages forming a double-headed structure.

of  $\epsilon_{mt}$  and  $\lambda_b$ .

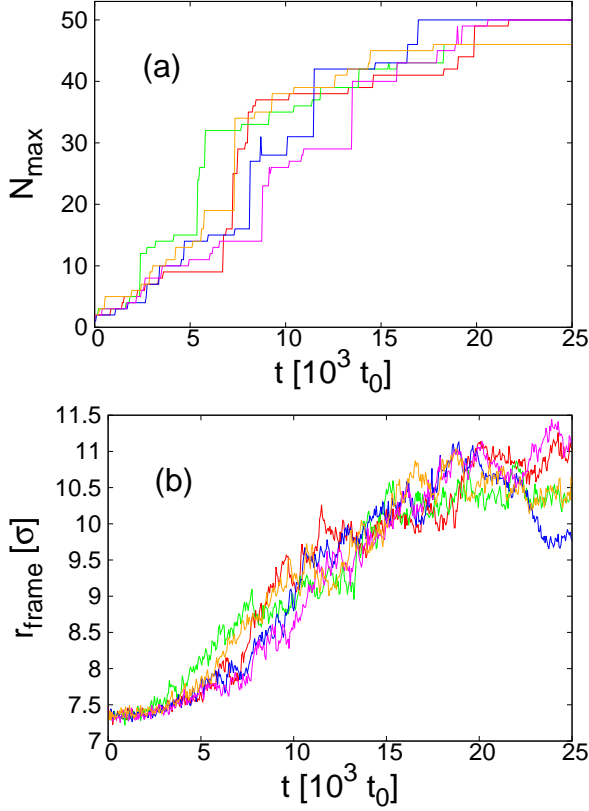
For the stronger attraction of the triskelia to the membrane,  $\epsilon_{mt} = 2k_B T$ , the rate of bud formation was similar all  $\lambda_b$  considered and the results were also similar for  $\epsilon_{mt} = k_B T$  with  $\lambda_b = 2\sqrt{3}k_B T$ . However, with the weaker attraction to the membrane,  $\epsilon_{mt} = k_B T$ , when the stiffness of the membrane was increased to  $\lambda_b = 4\sqrt{3}k_B T$ , the rate was significantly slower, although clear buds were formed. Increasing the membrane stiffness further to  $\lambda_b = 8\sqrt{3}k_B T$ , again with  $\epsilon_{mt} = k_B T$ , no clear buds were formed within  $2.5 \times 10^4 t_0$ , although caps on the membrane with some curvature were formed in some runs. The assembly of extended flat sheets



**Fig. 8** Average over 5 independent runs of  $r_{frame}$  against time for different parameters. Errorbars show the standard deviation of the data. They are only plotted intermittently and for some curves for clarity. Red:  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\lambda_b = 2\sqrt{3}k_B T$  for  $\epsilon_{mt} = k_B T$ ,  $\epsilon_{tt} = 5k_B T$  (with errorbars) and  $\epsilon_{mt} = k_B T$ ,  $\epsilon_{tt} = 10k_B T$  (without errorbars). Orange:  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\lambda_b = 2\sqrt{3}k_B T$ ,  $\epsilon_{mt} = 2k_B T$ ,  $\epsilon_{tt} = 5k_B T$ . Green:  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\lambda_b = 4\sqrt{3}k_B T$  for  $\epsilon_{mt} = k_B T$ ,  $\epsilon_{tt} = 5k_B T$  (with errorbars) and  $\epsilon_{mt} = 2k_B T$ ,  $\epsilon_{tt} = 10k_B T$  (without errorbars). Blue:  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\lambda_b = 8\sqrt{3}k_B T$  for  $\epsilon_{mt} = k_B T$ ,  $\epsilon_{tt} = 10k_B T$  (with errorbars) and  $\epsilon_{mt} = 2k_B T$ ,  $\epsilon_{tt} = 10k_B T$  (without errorbars). Magenta:  $\kappa = 3.2 \times 10^3 k_B T$ ,  $\lambda_b = 4\sqrt{3}k_B T$  for  $\epsilon_{mt} = 2k_B T$ ,  $\epsilon_{tt} = 10k_B T$  (with errorbars) and  $\epsilon_{mt} = k_B T$ ,  $\epsilon_{tt} = 5k_B T$  (without errorbars).

as in previous work<sup>15</sup> did not occur.

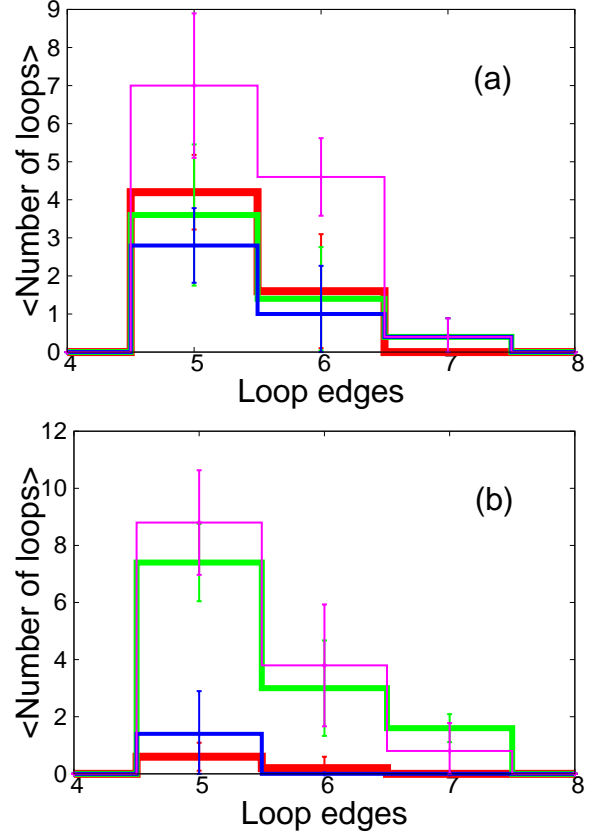
A typical pathway to bud formation was for multiple smaller clusters to form on the membrane, see for example



**Fig. 9** Individual runs for  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\lambda_b = 8\sqrt{3}k_B T$  for  $\epsilon_{mt} = 2k_B T$ ,  $\epsilon_{tt} = 5k_B T$ . (a) The population of the largest cluster in the system,  $N_{max}$ , against time. (b)  $r_{frame}$  against time. Curves of the same colour in the different plots show data for the same run. The blue curves show data for the run that is depicted in snapshots in Fig. 6.

the second snapshots in Fig. 6(a) and (b), and then coalesce, leading to a more rapid increase in the membrane curvature and bud formation, see for example the third and fourth snapshots in Fig. 6(a) and (b). For many runs, though not all, the footprint of this pathway could be seen by comparing the number of triskelia in the largest cluster in the system,  $N_{max}$ , and  $r_{frame}$  as a function of time. As may be seen by comparing Fig. 9(a) and Fig. 9(b) the most rapid increase in  $r_{frame}$  is correlated with a rapid increase in  $N_{max}$ , corresponding to smaller clusters joining together.

We finally, in Fig. 10, plot the distribution of the number of edges in the closed loops surrounding faces formed by assembled triskelia at the end of the simulation. It should be noted that this includes a contribution from assembly in the bulk as well as on the membrane, although this should be similar for all  $\epsilon_{mt}$  and  $\lambda_b$  and was small, as may be seen from the results for  $\epsilon_{mt} = k_B T$  in Fig. 10(b). Although large variation



**Fig. 10** Histograms of the number of closed loops surrounding faces formed by the triskelia with different numbers of edges at  $t = 2.5 \times 10^4 t_0$ . Averaged over 5 independent runs, errorbars show standard deviations. (a)  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\lambda_b = 2\sqrt{3}k_B T$  (b)  $\kappa = 0.8 \times 10^3 k_B T$ ,  $\lambda_b = 8\sqrt{3}k_B T$ . For both plots the different colours denote the same  $\epsilon_{tt}$  and  $\epsilon_{mt}$  values:  $\epsilon_{tt} = 5k_B T$ ,  $\epsilon_{mt} = k_B T$  (red);  $\epsilon_{tt} = 5k_B T$ ,  $\epsilon_{mt} = 2k_B T$  (green);  $\epsilon_{tt} = 10k_B T$ ,  $\epsilon_{mt} = k_B T$  (blue);  $\epsilon_{tt} = 10k_B T$ ,  $\epsilon_{mt} = 2k_B T$  (magenta). The thickness of the lines is varied for clarity.

was seen, generally more loops with 5 edges than loops with 6 edges were formed. For some parameters, 7-edge loops were formed.

## 5 Conclusions

We have introduced a new clathrin model using patchy beads that allows the inclusion of excluded volume and flexibility, as well as the interweaving of triskelia in assembled structures. Further, we have also described an approach to producing parameters for the model that will allow the assembly of similar structures to those seen in nature. Choosing the hexagonal barrel as a target structure, we employed our approach to find a parameter set. MC simulations using these parameters showed



that the triskelia could assemble a hexagonal barrel, as well as other structures observed in nature: the mini-coat and the tennis ball. Additionally, further structures were formed with different numbers of triskelia but also with 12 pentagonal and  $(N - 20)/2$  hexagonal faces, where  $N$  is the number of triskelia. The MC simulations were found to not be able to access all of the various local minima within one run. The mini-coat, hexagonal barrel, tennis ball, as well as truncated icosahedron structures were found to be mechanically stable for a range of triskelion stiffnesses, for sufficient attraction strengths.

Dynamical simulations of the assembly of the model triskelia with an attractive fluctuating membrane, employing a new membrane boundary condition, were performed with coupling to a coarse-grained solvent to include hydrodynamic interactions. For most parameters, the formation of buds by the assembly of the triskelia on the membrane surface was found. The buds were surrounded by cages with pentagonal, hexagonal and sometimes heptagonal faces. They often contained defects or holes and sometimes had lumpy, double-headed structures.

Our model takes into consideration key characteristics of clathrin, such as excluded volume, flexibility and binding site selectivity, whilst at the same time remaining computationally tractable. It is capable of reproducing the salient observed features of the protein: the assembly and stability of known structures and the formation of buds on a membrane. Whilst the smaller number of beads used in the current model is advantageous for simulation, its success suggests it could be interesting in further work to consider a similar model with a finer coarse-graining, that might be able to capture even more features. A similar approach might also be applied to some of the other proteins that attach to the membrane during budding<sup>33</sup>, and it could be very interesting to model their collaborative binding.

This work was supported by the Austrian Science Fund (FWF): M1367. Snapshots were created using VMD<sup>34</sup>. The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC).

## A Appendix

We present additional details of our model and methods. Many features of our model are similar to our previous work<sup>15</sup> and, correspondingly, parts of the descriptions in this appendix are very similar to parts of the supplemental material in ref.<sup>15</sup>. They are nonetheless reproduced here for the convenience of the reader.

### A.1 Triskelion-triskelion interactions

We first discuss the form of the triskelion-triskelion,  $tt$ , interactions, which have the same radial form as the triskelion-

membrane particle,  $mt$ , interactions. The potential form is similar to that used in earlier work<sup>19</sup>. For two different particles,  $i$  and  $j$ , separated by  $r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_j - \mathbf{r}_i|$ , where  $\mathbf{r}_i$  is position of particle  $i$ , the general form for both these types of interactions is,

$$U_{ij} = \gamma_{area} [U_{WCA}(r_{ij}) + \gamma_{att}\gamma_{orient}U_{att}(r_{ij})]$$

$$U_{WCA}(r) = \begin{cases} 4\epsilon \left[ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 + \frac{1}{4} \right] & \text{for } r < r_t, \\ 0 & \text{for } r \geq r_t, \end{cases}$$

$$U_{att}(r) = \begin{cases} -\epsilon & \text{for } r < r_t, \\ 4\epsilon \left[ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] & \text{for } r_t \leq r \leq r_s(1) \\ a(r - r_c)^2 + b(r - r_c)^3 & \text{for } r_s \leq r \leq r_c, \\ 0 & \text{for } r \geq r_c, \end{cases}$$

where  $r_t = 2^{1/6}\sigma$ ,  $r_s = (\frac{26}{7})^{1/6}\sigma$ ,  $r_c = \frac{67}{48}r_s$ ,  $a = -\frac{24192}{3211}\frac{\epsilon}{r_s^2}$  and  $b = -\frac{387072}{61009}\frac{\epsilon}{r_s^3}$ . The form of  $U_{att}(r)$  in the range  $r_s \leq r \leq r_c$  is a polynomial interpolation used to avoid a jump in the potential or its first derivative at the cut-off<sup>35</sup>. The energy scale,  $\epsilon$ , is set to  $\epsilon_{tt}$  for triskelion-triskelion interactions. The dimensionless factors  $\gamma_{area}$ ,  $\gamma_{att}$  and  $\gamma_{orient}$  take different forms for  $tt$  and  $mt$  interactions.

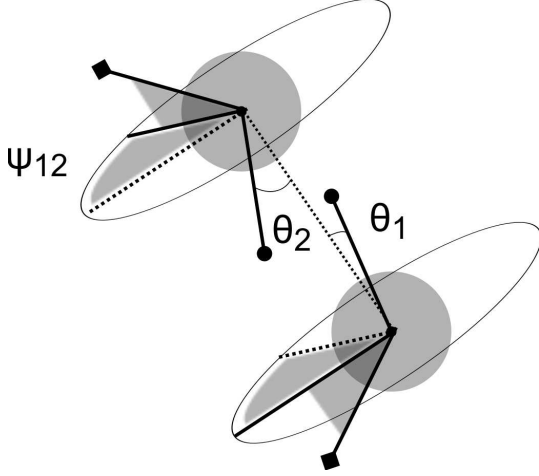
For  $tt$  interactions  $\gamma_{area} = 1$ . The patches on triskelion beads are given identities and only specific pairs are attractive, as detailed in the main text. For a pair of interacting triskelion beads from different triskelia, if the pair of closest patches are attractive then  $\gamma_{att} = 1$ , otherwise  $\gamma_{att} = 0$ .

The factor  $\gamma_{orient}$  allows the attractive part of the interaction to be made patchy. For  $tt$ -interactions the centres of the attractive patches are defined by unit vectors. For a given bead of type  $\alpha$ , the relative directions of these are given by the  $\{\mathbf{v}[\alpha, i]\}$  parameters. The width of the patches is determined by  $\gamma_{orient}$ , which is a product of functions of the form<sup>36</sup>:

$$F(\phi; \phi_a, \phi_b) = \begin{cases} 1 & \text{for } \phi \leq \phi_a, \\ \cos^2[(\pi/2)(\phi - \phi_a)/\phi_b] & \text{for } \phi_a \leq \phi \leq \phi_a + \phi_b, \\ 0 & \text{for } \phi \geq \phi_a + \phi_b. \end{cases} \quad (2)$$

$\gamma_{orient}(\hat{\mathbf{r}}_{ij}, \mathbf{\Omega}_i, \mathbf{\Omega}_j) = F(\theta_i; \theta_a, \theta_b) \times F(\theta_j; \theta_a, \theta_b) \times F(\psi_{ij}; 2\theta_a, 2\theta_b)$ , where  $\mathbf{\Omega}$  describes particle orientation.  $\theta_i$

is the angle between the interacting patch on particle  $i$  and  $\hat{\mathbf{r}}_{ij}$ , whilst  $\theta_j$  is between the patch of particle  $j$  and  $-\hat{\mathbf{r}}_{ij}$ .  $\psi_{ij}$  is the angle between the projections of the external torsional vectors of  $i$  and  $j$  onto the plane perpendicular to  $\hat{\mathbf{r}}_{ij}$ . The factor  $F(\psi_{ij})$  penalizes the twisting of interacting sub-units. We follow ref.<sup>19</sup> in choosing the range for this factor to be double that for the other ones. We choose  $\theta_a = 0.3$  and  $\theta_b = 0.2$ . The same geometry as for the external patchy interactions is also used for the internal interactions, see Fig. 11.



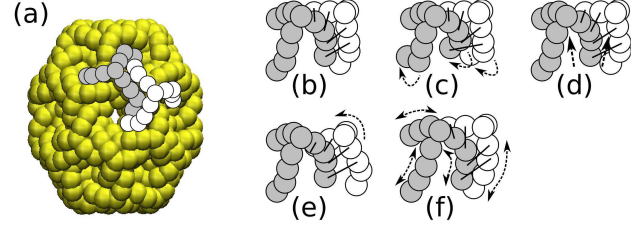
**Fig. 11** Sketch to illustrate the use of the unit vectors associated with bonded beads in a triskelion in calculating the internal interactions between those beads. The bending potential depends on the angles between the vectors terminating in circles, denoted bending vectors, and the bead-to-bead vector. The torsion potential depends on the angles between the projections of the two vectors terminating in diamonds, denoted internal torsional vectors, onto the plane perpendicular to the bead-to-bead vector. For a given bead of type  $\alpha$  the relative directions of the different vectors for internal interactions are defined by the  $\{\mathbf{u}[\alpha, i]\}$  parameters.

## A.2 Triskelion internal interactions

We next define the internal interactions between bonded beads within the same triskelion. Similarly to the external attractive patchy interactions, these are based on unit vectors associated with the bonded bead types. For a given bead of type  $\alpha$  their relative directions are given by the  $\{\mathbf{u}[\alpha, i]\}$  parameters. Fig. 11 depicts how the vectors are used to determine a set of angles between a pair of bonded beads. For these internal interactions, we denote the vectors whose angles to the bead-to-bead vector are considered as bending vectors, whilst the vectors whose rotations around it are considered are denoted internal torsional vectors. It should be noted that the torsional vectors for external and internal interactions are not the same.

A potential is then applied,  $U_{\text{internal}} = \frac{1}{2}k(r-d)^2 + \kappa(1 - \cos(\theta_1)) + \kappa(1 - \cos(\theta_2)) + \kappa(1 - \cos(\psi_{12}))$ , where  $r$  is the bead separation,  $\theta_1$  and  $\theta_2$  are angles between the bending vectors from the two beads respectively and the centre-to-centre vector.  $\psi_{12}$  is the angle between the projections of the two internal torsional vectors from the beads onto the plane perpendicular to the bead-to-bead vector.

## A.3 Triskelion parameter determination



**Fig. 12** Sketch of the process used to find a triskelion shape and set of interactions. (a) 36 triskelia are arranged with their centres on the vertices of a hexagonal barrel. Two neighbouring triskelia are highlighted in white and grey: (b) The pairs of beads that interact are indicated by thick connecting lines. Low-temperature Monte Carlo simulations are performed with a range moves: (c) A vector,  $\mathbf{u}^*$ , that determines the internal triskelion interactions is changed. Note that this changes the shape of all legs on all triskelia. Vectors defining the patches for external interactions,  $\mathbf{v}^*$  are also changed (not depicted). (d) The edge length,  $l_{\text{hex}}$  of the hexagonal barrel is varied, all triskelia receive a corresponding radial displacement. (e) Individual triskelia are rotated around their centre. (f) The separation of beads along the legs,  $d$ , is changed. In the latter stages of the simulation, triskelion centres are allowed to move freely.

We next give more details of the simulations used to determine triskelion parameters. 36 triskelia were placed with their centres on the vertices of a hexagonal barrel of edge length  $l_{\text{hex}}$ , although it should be noted that the hexagonal barrel structure cannot be formed from regular hexagons and pentagons<sup>4</sup> and so there is some ambiguity as to the exact positions. Initial values for  $\{\mathbf{v}^*[\alpha, i]\}$ ,  $\{\mathbf{u}^*[\alpha, i]\}$  and  $d$  were chosen by hand so that triskelia pairs interact in approximately the desired way. For the initial part of the simulation, triskelion centre positions remained fixed, except in updates of  $l_{\text{hex}}$  with a corresponding radial displacement of all centres. Further moves, summarised in Fig. 12, included varying  $d$  between  $0.5 \times 2^{1/6}\sigma$  and  $1.3 \times 2^{1/6}\sigma$ , as well as updates of  $\{\mathbf{v}^*[\alpha, i]\}$  and  $\{\mathbf{u}^*[\alpha, i]\}$ , and rigid-body rotations around triskelion centres. Due to the uncertainty about vertex positions, in the latter stages of the simulation, once the system energy was  $< 50\%$  of the possible minimum, the triskelion centres were allowed to move freely.

Vector identity		Bonding to		Vector components		
$\alpha$	$i$	$\beta$	$j$	$x$	$y$	$z$
$A$	1	$B$	1	0.95812	-0.28261	0.04633
	2	$B$	1	-0.68000	-0.68086	-0.27208
	3	$B$	1	-0.19950	0.96347	-0.17868
	4	<i>torsion</i>		0.19081	0.00000	-0.98163
$B$	1	$A$	1,2,3	-0.95812	0.28261	-0.04633
	2	$C$	1	0.80148	-0.32975	-0.49889
	3	<i>torsion</i>		-0.20924	0.11629	-0.97092
$C$	1	$B$	2	-0.80148	0.32975	0.49889
	2	$D$	1	0.73354	0.08292	-0.67456
	3	<i>torsion</i>		-0.55936	0.26125	-0.78668
$D$	1	$C$	2	-0.73354	-0.08292	0.67456
	2	$E$	1	0.14136	-0.48848	-0.86105
	3	<i>torsion</i>		-0.81931	0.22663	-0.52665
$E$	1	$D$	2	-0.14136	0.48848	0.86105
	2	<i>torsion</i>		-0.87421	0.34653	-0.34012

**Table 1** The vectors,  $\{\mathbf{u}[\alpha, i]\}$ , extracted from the parameter-finding simulations and employed in the internal interactions between bonded triskelion beads that determine the mechanical equilibrium shape in the free assembly simulations. The first two columns show the bead type,  $\alpha$ , and vector index,  $i$ . The second two columns show the bead type,  $\beta$ , and index,  $j$  corresponding to the other bending vector involved in the interaction, or alternatively show the label *torsion*, indicating the vector is used in the torsional component of interactions. Each bead type only has one torsional vector for internal interactions, which is used for all such interactions. Vector components are given to 5 decimal places.

Patch identity		Attracts		Vector components		
$\alpha$	$i$	$\beta$	$j$	$x$	$y$	$z$
$B$	1	$C$	1	0.49296	0.78900	-0.36669
	2	$D$	1	-0.31730	-0.43760	-0.84132
	3	<i>torsion</i>		0.84037	-0.54084	-0.03577
$C$	1	$B$	1	0.52052	0.79423	0.31345
	2	$E$	1	-0.40401	-0.09107	-0.91021
	3	<i>torsion</i>		-0.84655	0.53135	0.03203
$D$	1	$B$	2	0.82034	-0.13387	0.55599
	2	<i>torsion</i>		-0.14006	-0.97990	-0.14203
$E$	1	$C$	2	0.62525	-0.39778	0.67144
	2	<i>torsion</i>		0.24275	0.92872	0.28026

**Table 2** The vectors,  $\{\mathbf{v}[\alpha, i]\}$ , extracted from the parameter-finding simulations and employed in the external interactions between triskelion beads from different triskelia in the free assembly simulations. The first two columns show the bead type,  $\alpha$ , and patch index,  $i$ . The second two columns show the bead type,  $\beta$ , and index,  $j$  corresponding to the other patch that the first patch attracts, or alternatively show the label *torsion*, indicating the vector is used in the torsional component of interactions. Each bead type only has one torsional vector for external interactions, which is used for all such interactions. Vector components are given to 5 decimal places.

Our triskelion model is unchanged by the application of an arbitrary rotation to all vectors associated with a bead type: since the internal interactions between different beads fix their relative orientations, only the directions of the vectors for a particular type relative to each other are important. It is thus necessary to constrain some vectors during interaction-finding

simulations. Without loss of freedom, for each bead type  $B - E$ , we chose to fix the bending vector used for the interaction with the previous bead in the leg. Additionally, the internal torsional vector for type  $A$ , which lies on its axis of symmetry, was fixed, and the three bending vectors of  $A$  for interactions with  $B$  beads were constrained to not rotate around the

symmetry axis. The internal interaction vectors for bead type  $A$  were such that there was a threefold rotational symmetry axis through the centre.

If such constraints were not applied, vectors would be able to explore all directions and their average would become undefined. Furthermore, even with the constraints already outlined, there is a large degeneracy in possible torsional vectors, both for internal and external interactions. Therefore, internal torsional vectors, for all bead types except  $A$ , were adjusted for each configuration to make them as close to perpendicular to the bending vectors as possible, without altering the resulting triskelion. Furthermore, the external torsional vector for bead type  $B$  was taken to be  $\mathbf{v}^*[B, 2] \times \mathbf{v}^*[B, 1] / |\mathbf{v}^*[B, 2] \times \mathbf{v}^*[B, 1]|$ . For the other external torsional vectors, a minimum angle between attractive patch vectors and their corresponding torsional vector was imposed.

Not all interaction-finding simulations were found to converge to a low-energy minimum with all patches interacting as desired. Therefore, an intermediate configuration from a successful simulation, for which all patches were interacting, was chosen as a starting point for further simulations. These were run with 3 repeats each at  $k_B T = 10^{-2}, 10^{-3}$  and  $10^{-4} \epsilon_{tt}$ . At a given  $k_B T$ , all simulations were found to converge to states fluctuating around the same energy value. As expected, these energy values, and the size of the fluctuations, both relative to  $\epsilon_{tt}$ , decreased with decreasing  $k_B T$ . All were within a range of  $\approx \epsilon_{tt}$  around  $-310 \epsilon_{tt}$ , compared to a possible minimum of  $-324 \epsilon_{tt}$  if each attractive patch on each triskelion interacted with one other patch and all these interactions were minimised.

Sets of parameters for free assembly simulations were extracted as thermal averages, taken after the simulation had relaxed to the minimum,  $d = \langle d^* \rangle$ ,  $\{\mathbf{v}[\alpha, i]\} = \{\langle \mathbf{v}^*[\alpha, i] \rangle / |\langle \mathbf{v}^*[\alpha, i] \rangle|\}$  and  $\{\mathbf{u}[\alpha, i]\} = \{\langle \mathbf{u}^*[\alpha, i] \rangle / |\langle \mathbf{u}^*[\alpha, i] \rangle|\}$ . The parameters from all simulations at different  $k_B T$  were very similar. To test the size of the differences between the different parameter sets, triskelia defined by each were placed, in their mechanical equilibrium configurations, with the same centre position and with the centre symmetry vector and the direction of the first leg aligned. The maximum distance between corresponding beads in different triskelia was  $7 \times 10^{-3} \sigma$  and the largest angle between patches was  $4\pi \times 10^{-3}$ . The largest angle between torsion vectors was larger,  $6\pi \times 10^{-2}$ , due to the larger freedom in choosing these. Given the close similarity, we simply picked one set and verified that these parameters did produce triskelia that can form a hexagonal barrel with energy  $\approx -310 \epsilon_{tt}$  for low  $k_B T$ .

We give the parameters used in our free assembly simulations. The equilibrium separation between internally bonded triskelion beads was  $d = 0.81597 \sigma$ , to 5 decimal places. The vector parameters used for defining internal,  $\{\mathbf{u}[\alpha, i]\}$ , and

external,  $\{\mathbf{v}[\alpha, i]\}$ , interactions are given in Tables 1 and 2 respectively.

#### A.4 Membrane Model

We next discuss our membrane model. For the interactions between membrane particles, we use smooth potentials that are appropriate for molecular dynamics<sup>28</sup>. Bonded membrane particles interact via

$$U_{bond}(r_{ij}) = \begin{cases} 0 & \text{for } r_{ij} \leq 1.15\sigma, \\ (80k_B T) \exp[1/(1.15\sigma - r_{ij})]/(1.33\sigma - r_{ij}) & \text{for } 1.15\sigma < r_{ij} < 1.33\sigma, \\ \infty & \text{for } r_{ij} \geq 1.33\sigma, \end{cases} \quad (3)$$

with  $r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_j - \mathbf{r}_i|$ , where  $\mathbf{r}_i$  is position of particle  $i$ . All pairs of membrane particles experience an excluded volume potential

$$U_{EV}(r_{ij}) = \begin{cases} \infty & \text{for } r_{ij} \leq 0.67\sigma, \\ (80k_B T) \exp[1/(r_{ij} - 0.85\sigma)]/(r_{ij} - 0.67\sigma) & \text{for } 0.67\sigma < r_{ij} < 0.85\sigma, \\ 0 & \text{for } r_{ij} \geq 0.85\sigma. \end{cases} \quad (4)$$

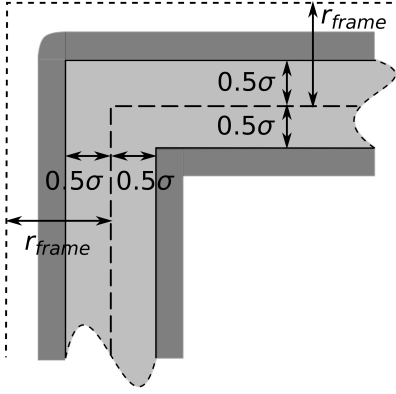
The minimum distance between any two membrane particles is  $0.67\sigma$  and the maximum bond length is  $1.33\sigma$ .

Since we employ an SRD solvent, which requires the simulation box to be divisible into a regular grid of cells, the approach of using box-rescaling for simulating a tensionless membrane from our previous work without solvent<sup>15</sup> is not suitable. We therefore simulate the membrane by restricting a number,  $N_{border}$ , of particles, denoted border particles, to a frame with a confining potential. These form the edge of the membrane. Other, bulk, particles do not experience the confining potential. Furthermore, bonds between membrane particles are also of border and bulk types, where the border bonds are always between two border particles and always form a single closed ring. The bonding potential for both bond types is the same.

The position of the frame is given by a variable  $r_{frame}$  that defines how far in from the edges of the simulation box it is, see Fig. 13. In the direction out of the plane of the frame a flat-potential region extends for  $4\sigma$ . In the plane it extends  $0.5\sigma$  inwards and outwards from the frame position, see Fig. 13.

In the flat-potential region, border membrane particles have an energy of  $E_{frame}$ . The confining potential for the border membrane particles is of the same form, range and strength as





**Fig. 13** Sketch of a section of the frame in a simulation box, showing a cut in the plane of the frame. The outer, short-dashed line shows the periodic boundaries of the simulation box. The centre of the frame, shown by the long-dashed line is located a distance  $r_{frame}$  inwards from this. There is a region where border membrane particles experience a flat potential, extending  $0.5\sigma$  in either direction from this, shown by the light grey area. In the dark grey area, the border membrane particles experience a confining potential that diverges at the edge further from the flat region. In the direction out of the plane (not shown) the flat-potential region extends  $4\sigma$ .

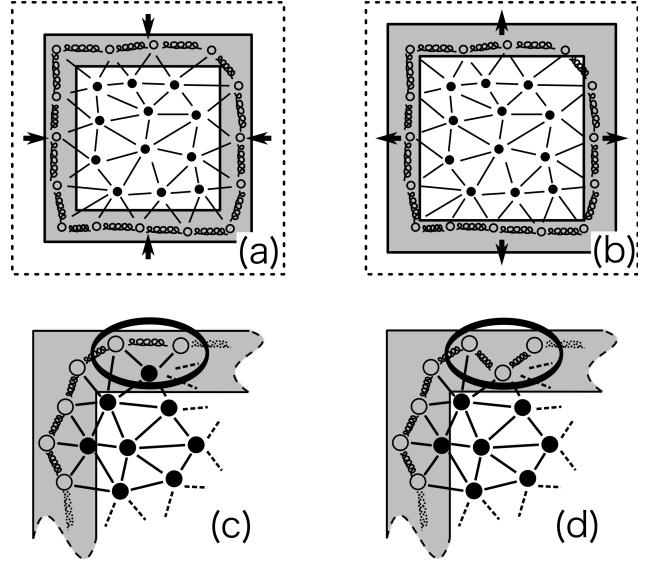
the excluded volume between membrane particles,

$$U_{confine}(r) = \begin{cases} E_{frame} + (80k_B T) \exp[-1/r]/(0.18\sigma - r) & \text{for } 0 < r < 0.18\sigma, \\ \infty & \text{for } r \geq 0.18\sigma, \end{cases} \quad (5)$$

where  $r$  is the distance of the border membrane particle from the closest point within the flat-potential region.

As summarised in Fig. 14, we employ two types of MC moves to allow the membrane to expand and contract: those that convert border particles to bulk particle and vice versa, as well as those that change  $r_{frame}$ , moving the frame inwards or outwards from the simulation box centre. The total number of bonds and triangles,  $N_{tri}$ , in the membrane may thus vary. Maximum and minimum values of  $r_{frame}$  are imposed such that the confining potentials from opposite sides of the frame do not overlap, and also such that the confining potential does not extend across the periodic boundaries.

The value of  $E_{frame}$  controls  $r_{frame}$  and  $N_{border}$ : if it is made very high the membrane will try to minimise  $N_{border}$ , and consequently  $r_{frame}$  will increase as the membrane is squeezed, forcing it to extend out of the plane. On the other hand, if  $E_{frame}$  is made very low,  $N_{border}$  will become large and  $r_{frame}$  small, and the central region of the membrane will be stretched into a flat configuration. To determine  $E_{frame}$



**Fig. 14** Sketch of the membrane boundary conditions used for dynamics simulations. Membrane particles are of two types: border (open circles) and bulk (filled circles), as are membrane bonds: border (loopy lines) and bulk (straight lines). In the shaded region, border particles experience a confining potential with a flat minimum region with energy  $E_{frame}$ , bulk particles feel no potential. Every so often, Monte Carlo updates of two types are performed: (a), (b) The position of the frame within the simulation box (dotted line) is shifted inwards or outwards. (c), (d) Bulk particles are converted to border particles, see highlighted area, or vice versa. Only bulk particles that are bonded to two current border particles, which are themselves bonded, may be converted. A corresponding bond creation (deletion), along with a conversion of existing bonds from bulk to border (border to bulk) ensures that each border particle has two, and only two, border bonds to two other border particles, and that the bonds between all the border particles form a closed ring. A double bond between the same two particles may not be created and a border bond may only be deleted if the particle that thus becomes part of the border was originally a bulk particle.

for a particular membrane stiffness, we simulated membranes with a range of  $E_{frame}$  and compared the ratio of the area to the area projected onto the plane of the frame,  $A/A_{proj}$ , to the value measured for a tensionless membrane simulated with box-rescaling<sup>15</sup>. To minimise the effect of the frame, we only considered the part of the membrane with a projection falling within the central 10% of the area defined by the frame.

A unit normal vector is associated with each membrane triangle. Each bulk bond forms the side of two different neighboring triangles. Membrane fluidity is included using MC moves that attempt to remove a given bulk bond and create a new one between the two vertices of its neighboring triangles that were not connected by the original. During dynamical

simulations, a number of moves equal to the number of bulk bonds are performed every  $0.1t_0$ . By simulating a Poiseuille flow for a two-dimensional membrane<sup>28</sup>, we estimate the resulting viscosity of the membrane to be  $35.1 \pm 0.1m/t_0$ . During the bond-flipping procedure, the direction of the normals is always maintained such that, if the membrane were in a flat configuration, all normals would be aligned.

The bending stiffness of the membrane is controlled by including a potential  $U_{bend} = \lambda_b(1 - \mathbf{n}_i \cdot \mathbf{n}_j)$  for each bond, where  $\mathbf{n}_i$  and  $\mathbf{n}_j$  are the unit normal vectors of the two triangles neighboring the bond and  $\lambda_b$  is an energy. The total membrane area,  $A$ , is constrained with a harmonic potential,  $U_{area} = (k_B T)(A - A_0)^2$ , where  $A_0 = (\sqrt{3}/4)l^2 N_{tri}$ , in the Hamiltonian. Additionally, a bending potential of the same form is applied to triangles with a border bond as one of their edges, where the unit normal of the triangle is compared to a unit normal to the frame-plane.

In our dynamical simulations, a series of MC moves, each changing  $r_{frame}$  by  $\Delta r_{frame}$  or  $N_{border}$  by  $\pm 1$ , were performed every  $0.1t_0$ . On average there were  $10^3$  attempted changes to  $r_{frame}$  and  $10N_{mem}$  attempted moves to change  $N_{border}$ . The acceptance ratio for the  $r_{frame}$ -moves showed some dependence on the simulation parameters and also changed somewhat during the course of a simulation, for example as a bud was formed. It was however found to always be roughly in the range  $0.3 - 0.4$ . The acceptance ratio for the  $N_{border}$ -moves depended on the membrane stiffness, ranging from  $\approx 5 \times 10^{-2}$  for  $\lambda_b = 8\sqrt{3}k_B T$  to  $\approx 7 \times 10^{-2}$  for  $\lambda_b = 2\sqrt{3}k_B T$ . A similar, though weaker, dependence was seen for the acceptance ratio of bond flips, which had a value of  $\approx 2 \times 10^{-2}$ , similar to that seen in previous work with the same membrane model<sup>28</sup>.

The rate of  $N_{border}$ -moves chosen means there were  $\approx 600$  such moves accepted every  $0.1t_0$ . Similarly, there were  $\approx 300$   $r_{frame}$ -moves accepted.  $\Delta r_{frame}$  values are chosen uniformly in the range  $-0.05\sigma < \Delta r_{frame} < 0.05\sigma$  so if these  $\approx 300$  shifts were an unbiased random walk,  $r_{frame}$  would explore a range of  $\approx 0.2\sigma$ . Given that the values of  $N_{border}$  and  $r_{frame}$  changed by at most  $\approx 80$  and  $\approx 7\sigma$  respectively over the course of a  $2.5 \times 10^4 t_0$  simulation, the rate of  $N_{border}$ - and  $r_{frame}$ -moves are sufficiently high that they will adjust the system to the local minimum every  $0.1t_0$  and their exact values will not affect the results.

We next discuss the triskelion-membrane particle interactions, the form of which is given in Eq. 1. For these interactions, the energy scale,  $\epsilon$ , is set to  $\epsilon_{mt}$ . Since the membrane bonds have a relatively broad, flat minimum, the membrane particles would tend to be locally compressed when an attractive triskelion bead is close. For  $mt$  interactions the  $\gamma_{area}$  factor is used to counter-act this by making the interaction proportional to the area that the membrane particle represents:  $\gamma_{area} = A_{neigh}/(N_{neigh}A_{tri})$ , where  $N_{neigh}$  is the

total number of triangles that have the membrane particle as a vertex,  $A_{neigh}$  is their total area and  $A_{tri} = A_0/N_{tri}$ .

For  $mt$  interactions,  $\gamma_{att} = 0$  for triskelion bead types  $A - D$ . For bead type  $E$  it is used to make only one side of the membrane attract the triskelion beads: it takes a value of 1 if the bead is “above” the membrane and 0 if it is “below”. A triskelion bead is determined to be “above” or “below” by finding the closest point on the membrane. If the normal of the triangle enclosing the closest point makes an angle of less than  $\pi/2$  with the vector from the closest point to bead then the bead is “above” the membrane, otherwise it is “below”. For  $mt$  interactions  $\gamma_{orient} = 1$  since these are chosen not to be patchy.

Since the attractive interaction between triskelia and the membrane depends on which side of the membrane they interact with, a discontinuity in the potential would arise if triskelia could move from one side of the membrane to the other whilst remaining within the interaction range. To avoid this, triskelia experience an excluded volume potential around the frame. So that the total assembly volume available to the triskelia remains approximately constant, the width of the excluded volume region around the frame is rescaled as it moves so its volume is not changed. This excluded region does not generally extend across the entire simulation box.

## A.5 SRD

We finally discuss our choice of SRD parameters. The side of the SRD cells was chosen to be equal to  $\sigma$ . Both membrane particles and the beads forming triskelia were given masses of  $5m$  and coupled to the solvent via the collision step<sup>29</sup>. The rotational degrees of freedom of the beads are thus not coupled directly to the solvent but, given the relatively stiff triskelia we simulate, these relax rapidly anyway. In mechanical equilibrium the separation between the central triskelion bead and the final bead of a leg is about  $2.8\sigma$  and that between the final beads of different legs is about  $3.4\sigma$ . A triskelion thus spans many SRD cells and so its rotation as a whole is coupled to the solvent. We chose to have a number density of 5 SRD particles per cell and performed collisions every  $0.1t_0$  with a rotation angle of  $\frac{\pi}{2}$ , giving a viscosity of  $\eta = 2.5m/\sigma t_0$ <sup>37</sup>. We apply a momentum-conserving cell-level thermostat to the SRD fluid<sup>29</sup>.

## References

1. A. Fotin, Y. Cheng, P. Sliz, N. Grigorieff, S. C. Harrison, T. Kirchhausen and T. Walz, *Nature*, 2004, **432**, 573–579.
2. F. Brodsky, C. Chen, C. Knuehl, M. Towler and D. Wakeham, *Annu. Rev. Cell Dev. Biol.*, 2001, **17**, 517–568.
3. J. Heuser, *J. Cell Biol.*, 1980, **84**, 560–583.
4. S. Schein and M. Sands-Kidner, *Biophys. J.*, 2008, **94**, 958–976.
5. S. Zaremba and J. Keen, *J. Cell Biol.*, 1983, **97**, 1339–1347.

- 6 B. Pearse and R. Crowther, *Annu. Rev. Biophys. Biophys. Chem.*, 1987, **16**, 49–68.
- 7 A. Fotin, T. Kirchhausen, N. Grigorieff, S. C. Harrison, T. Walz and Y. Cheng, *J. Struct. Biol.*, 2006, **156**, 453–460.
- 8 Y. Cheng, W. Boll, T. Kirchhausen, S. C. Harrison and T. Walz, *J. Mol. Biol.*, 2007, **365**, 892–899.
- 9 W. den Otter, M. Renes and W. Briels, *J. Phys.: Cond. Matt.*, 2010, **22**, 104103.
- 10 W. den Otter, M. Renes and W. Briels, *Biophys. J.*, 2010, **99**, 1231–1238.
- 11 A. J. Jin and R. Nossal, *Biophys. J.*, 2000, **78**, 1183–1194.
- 12 M. L. Ferguson, K. Prasad, H. Boukari, D. L. Sackett, S. Krueger, E. M. Lafer and R. Nossal, *Biophys. J.*, 2008, **95**, 1945–1955.
- 13 T. Kirchhausen, *Annu. Rev. Biochem.*, 2000, **69**, 699–727.
- 14 J. A. Ybe, F. M. Brodsky, K. Hofmann, K. Lin, S.-H. Liu, L. Chen, T. N. Earnest, R. J. Fletterick, P. K. Hwang *et al.*, *Nature*, 1999, **399**, 371–374.
- 15 R. Matthews and C. N. Likos, *Phys. Rev. Lett.*, 2012, **109**, 178302.
- 16 D. Rapaport, *Physics Procedia*, 2012, **34**, 114–118.
- 17 B. Reynwar, G. Illya, V. Harmandaris, M. Müller, K. Kremer and M. Deserno, *Nature*, 2007, **447**, 461–464.
- 18 M. F. Hagan, *Adv. Chem. Phys. (arXiv:1301.1657)*, 2013, **155**, (in press).
- 19 A. Wilber, J. Doye, A. Louis and A. Lewis, *J. Chem. Phys.*, 2009, **131**, 175102.
- 20 D. Frenkel and B. Smit, *Understanding Molecular Simulation: from Algorithms to Applications*, Academic Press, London, 2002.
- 21 B. Chen and J. Siepmann, *J. Phys. Chem. B*, 2001, **105**, 11275–11282.
- 22 T. Vlught, M. Martin, B. Smit, J. Siepmann and R. Krishna, *Mol. Phys.*, 1998, **94**, 727–733.
- 23 A. Bhattacharyay and A. Troisi, *Chem. Phys. Lett.*, 2008, **458**, 210–213.
- 24 B. Mehlig, D. Heermann and B. Forrest, *Phys. Rev. B*, 1992, **45**, 679–685.
- 25 R. Faller, Q. Yan and J. De Pablo, *J. Chem. Phys.*, 2002, **116**, 5419.
- 26 J. Aronovitz and D. Nelson, *Journal de Physique*, 1986, **47**, 1445–1456.
- 27 L. Weinberg, *IEEE Trans. Circuit Theory*, 1966, **13**, 142–148.
- 28 H. Noguchi and G. Gompper, *Phys. Rev. E*, 2005, **72**, 11901.
- 29 G. Gompper, T. Ihle, D. Kroll and R. Winkler, *Advanced Computer Simulation Approaches for Soft Matter Sciences III*, 2009, 1–87.
- 30 N. Kikuchi, A. Gent and J. Yeomans, *Eur. Phys. J. E*, 2002, **9**, 63–66.
- 31 C. W. Harland, M. J. Bradley and R. Parthasarathy, *Proc. Nat. Acad. Sci. USA*, 2010, **107**, 19146–19150.
- 32 T.-G. Iversen, G. Skretting, B. van Deurs and K. Sandvig, *Proc. Nat. Acad. Sci. USA*, 2003, **100**, 5175–5180.
- 33 J. Liu, Y. Sun, D. Drubin and G. Oster, *PLoS Biology*, 2009, **7**, e1000204.
- 34 W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33–38.
- 35 P. Bordat, D. Reith and F. Müller-Plathe, *J. Chem. Phys.*, 2001, **115**, 8978.
- 36 W. Miller and A. Cacciuto, *Phys. Rev. E*, 2009, **80**, 021404.
- 37 N. Kikuchi, C. Pooley, J. Ryder and J. Yeomans, *J. Chem. Phys.*, 2003, **119**, 6388.